PLoS ONE, 2017, vol.12, N10

Structure of 311 service requests as a signature of urban location

Wang L., Qian C., Kats P., Kontokosta C., Sobolevsky S.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 Wang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. While urban systems demonstrate high spatial heterogeneity, many urban planning, economic and political decisions heavily rely on a deep understanding of local neighborhood contexts. We show that the structure of 311 Service Requests enables one possible way of building a unique signature of the local urban context, thus being able to serve as a low-cost decision support tool for urban stakeholders. Considering examples of New York City, Boston and Chicago, we demonstrate how 311 Service Requests recorded and categorized by type in each neighborhood can be utilized to generate a meaningful classification of locations across the city, based on distinctive socioeconomic profiles. Moreover, the 311-based classification of urban neighborhoods can present sufficient information to model various socioeconomic features. Finally, we show that these characteristics are capable of predicting future trends in comparative local real estate prices. We demonstrate 311 Service Requests data can be used to monitor and predict socioeconomic performance of urban neighborhoods, allowing urban stakeholders to quantify the impacts of their interventions.

http://dx.doi.org/10.1371/journal.pone.0186314

References

[1] Batty M. The size, scale, and shape of cities. science. 2008; 319(5864):769–771. https://doi.org/10.1126/science.1151419 PMID: 18258906

[2] Bettencourt LM, Lobo J, Strumsky D, West GB. Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. PloS one. 2010; 5(11):e13541. https://doi.org/10.1371/journal.pone.0013541 PMID: 21085659

[3] Bettencourt LM. The origins of scaling in cities. science. 2013; 340(6139):1438–1441. https://doi.org/10.1126/science.1235823 PMID: 23788793

[4] Arcaute E, Youn H, Ferguson P, Hatna E, Batty M, Johansson A. City boundaries and the universality of scaling laws; 2013.

[5] Kontokosta CE. The quantified community and neighborhood labs: A framework for computational urban science and civic technology innovation. Journal of Urban Technology. 2016; 23(4):67–84. https://doi.org/10.1080/10630732.2016.1177260

[6] Rokach L, Maimon O. The Data Mining and Knowledge Discovery Handbook: A Complete Guide for Researchers and Practitioners. New York: Springer; 2005.

[7] Townsend AM. Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia. W. W. Norton & Company; 2013.

[8] Powell LM, Slater S, Mirtcheva D, Bao Y, Chaloupka FJ. Food store availability and neighborhood characteristics in the United States. Preventive medicine. 2007; 44(3):189–195. https://doi.org/10.1016/j.ypmed.2006.08.008 PMID: 16997358

[9] Bettencourt LM, Lobo J, Helbing D, Kühnert C, West GB. Growth, innovation, scaling, and the pace of life in cities. Proceedings of the national academy of sciences. 2007; 104(17):7301–7306. https://doi.org/10.1073/pnas.0610172104

[10] Albeverio S, Andrey D, Giordano P, Vancheri A. The dynamics of complex urban systems: An interdisciplinary approach. Springer; 2007.

[11] Sobolevsky S, Sitko I, Grauwin S, Combes RTd, Hawelka B, Arias JM, et al. Mining urban performance: Scale-independent classification of cities based on individual economic transactions. arXiv preprint arXiv:14054301. 2014;.

[12] Sobolevsky S, Bojic I, Belyi A, Sitko I, Hawelka B, Arias JM, et al. Scaling of city attractiveness for foreign visitors through big data of human economical and social media activity. In: Big Data (BigData Congress), 2015 IEEE International Congress on. IEEE; 2015. p. 600–607.

[13] Nelder JA, Baker RJ. Generalized linear models. Wiley Online Library; 1972.

[14] Bolstad BM, Irizarry RA, \AAstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003; 19(2):185–193. https://doi.org/10.1093/bioinformatics/19.2.185 PMID: 12538238

[15] MacQueen J, others. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. vol. 1. Oakland, CA, USA.; 1967. p. 281–297.

[16] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics. 1987; 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[17] Allwinkle S, Cruickshank P. Creating smart-er cities: An overview. Journal of urban technology. 2011; 18(2):1–16. https://doi.org/10.1080/10630732.2011.601103

[18] Batty M. Smart cities, big data. SAGE Publications Sage UK: London, England; 2012.

[19] Bettencourt LM. The uses of big data in cities. Big Data. 2014; 2(1):12–22. https://doi.org/10.1089/big.2013.0042 PMID: 27447307

[20] Girardin F, Calabrese F, Dal Fiore F, Ratti C, Blat J. Digital footprinting: Uncovering tourists with user-generated content. IEEE Pervasive computing. 2008; 7(4). https://doi.org/10.1109/MPRV.2008.71

[21] Gonzalez MC, Hidalgo CA, Barabasi AL. Understanding individual human mobility patterns. arXiv preprint arXiv:08061256. 2008;.

[22] Quercia D, Lathia N, Calabrese F, Di Lorenzo G, Crowcroft J. Recommending social events from mobile phone location data. In: Data Mining (ICDM), 2010 IEEE 10th International Conference on. IEEE; 2010. p. 971–976.

[23] Amini A, Kung K, Kang C, Sobolevsky S, Ratti C. The impact of social segregation on human mobility in developing and industrialized regions. EPJ Data Science. 2014; 3(1):6. https://doi.org/10.1140/epjds31

[24] Reades J, Calabrese F, Sevtsuk A, Ratti C. Cellular census: Explorations in urban data collection. IEEE Pervasive computing. 2007; 6(3). https://doi.org/10.1109/MPRV.2007.53

[25] Kontokosta CE, Johnson N. Urban phenology: Toward a real-time census of the city using Wi-Fi data. Computers, Environment and Urban Systems. 2016; 64:144–153. https://doi.org/10.1016/j.compenvurbsys.2017.01.011

[26] Santi P, Resta G, Szell M, Sobolevsky S, Strogatz SH, Ratti C. Quantifying the benefits of vehicle pooling with shareability networks. Proceedings of the National Academy of Sciences. 2014; 111 (37):13290–13294. https://doi.org/10.1073/pnas.1403657111

[27] Sobolevsky S, Sitko I, des Combes RT, Hawelka B, Arias JM, Ratti C. Cities through the prism of people's spending behavior. PloS one. 2016; 11(2):e0146291. https://doi.org/10.1371/journal.pone.0146291 PMID: 26849218

[28] Shen S, Sam AG, Jones E. Credit Card Indebtedness and Psychological Well-Being Over Time: Empirical Evidence from a Household Survey. Journal of Consumer Affairs. 2014; 48(3):431–456. https://doi.org/10.1111/joca.12047

[29] Scholnick B, Massoud N, Saunders A. The impact of wealth on financial mistakes: Evidence from credit card non-payment. Journal of financial Stability. 2013; 9(1):26–37. https://doi.org/10.1016/j.jfs.2012.11.005

[30] Boeschoten WC. Cash management, payment patterns and the demand for money. De economist. 1998; 146(1):117–142. https://doi.org/10.1023/A:1003258026314

[31] Bounie D, Francois A. Cash, Check or Bank Card? The Effects of Transaction Characteristics on the Use of Payment Instruments. Rochester, NY: Social Science Research Network; 2006. ID 891791.

[32] Hayhoe CR, Leach LJ, Turner PR, Bruin MJ, Lawrence FC. Differences in spending habits and credit use of college students. Journal of Consumer Affairs. 2000; 34(1):113–133. https://doi.org/10.1111/j.1745-6606.2000.tb00087.x

[33] Bagchi M, White PR. The potential of public transport smart card data. Transport Policy. 2005; 12 (5):464–474. https://doi.org/10.1016/j.tranpol.2005.06.008

[34] Chan PK, Fan W, Prodromidis AL, Stolfo SJ. Distributed data mining in credit card fraud detection. IEEE Intelligent Systems and Their Applications. 1999; 14(6):67–74. https://doi.org/10.1109/5254.809570

[35] Rysman M. An empirical analysis of payment card usage. The Journal of Industrial Economics. 2007; 55(1):1–36. https://doi.org/10.1111/j.1467-6451.2007.00301.x

[36] Szell M, Grauwin S, Ratti C. Contraction of online response to major events. PloS one. 2014; 9(2): e89052. https://doi.org/10.1371/journal.pone.0089052 PMID: 24586499

[37] Frank MR, Mitchell L, Dodds PS, Danforth CM. Happiness and the patterns of life: A study of geolocated tweets. arXiv preprint arXiv:13041296. 2013;.

[38] Hawelka B, Sitko I, Beinat E, Sobolevsky S, Kazakopoulos P, Ratti C. Geo-located Twitter as proxy for global mobility patterns. Cartography and Geographic Information Science. 2014; 41(3):260–271. https://doi.org/10.1080/15230406.2014.890072 PMID: 27019645

[39] Paldino S, Bojic I, Sobolevsky S, Ratti C, González MC. Urban magnetism through the lens of geotagged photography. EPJ Data Science. 2015; 4(1):5. https://doi.org/10.1140/epjds/s13688-015-0043-3

[40] Lenormand M, Louail T, Cantú-Ros OG, Picornell M, Herranz R, Arias JM, et al. Influence of sociodemographic characteristics on human mobility. arXiv preprint arXiv:14117895. 2014;.

[41] Louail T, Lenormand M, Ros OGC, Picornell M, Herranz R, Frias-Martinez E, et al. From mobile phone data to the spatial structure of cities. Scientific reports. 2014; 4. https://doi.org/10.1038/srep05276 PMID: 24923248

[42] Ratti C, Sobolevsky S, Calabrese F, Andris C, Reades J, Martino M, et al. Redrawing the map of Great Britain from a network of human interactions. PloS one. 2010; 5(12):e14248. https://doi.org/10.1371/journal.pone.0014248 PMID: 21170390

[43] Sobolevsky S, Szell M, Campari R, Couronné T, Smoreda Z, Ratti C. Delineating geographical regions with networks of human interactions in an extensive set of countries. PloS one. 2013; 8(12):e81707. https://doi.org/10.1371/journal.pone.0081707 PMID: 24367490

[44] Grauwin S, Sobolevsky S, Moritz S, Gódor I, Ratti C. Towards a comparative science of cities: Using mobile traffic records in new york, london, and hong kong. In: Computational approaches for urban environments. Springer; 2015. p. 363–387.

[45] Pei T, Sobolevsky S, Ratti C, Shaw SL, Li T, Zhou C. A new insight into land use classification based on aggregated mobile phone data. International Journal of Geographical Information Science. 2014; 28 (9):1988–2007. https://doi.org/10.1080/13658816.2014.913794

[46] Sobolevsky S, Sitko I, Des Combes RT, Hawelka B, Arias JM, Ratti C. Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. the case of residents and foreign visitors in spain. In: Big Data (BigData Congress), 2014 IEEE International Congress on. IEEE; 2014. p. 136–143.

[47] Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C. A Tale of Many Cities: Universal Patterns in Human Urban Mobility. PLOS ONE. 2012; 7(5):e37027. https://doi.org/10.1371/journal.pone.0037027 PMID: 22666339

[48] Kung KS, Greco K, Sobolevsky S, Ratti C. Exploring universal patterns in human home-work commuting from mobile phone data. PloS one. 2014; 9(6):e96180. https://doi.org/10.1371/journal.pone.0096180 PMID: 24933264

[49] Lathia N, Quercia D, Crowcroft J. The hidden image of the city: sensing community well-being from urban mobility. Pervasive computing. 2012; p. 91–98.

[50] Sobolevsky S, Massaro E, Bojic I, Arias JM, Ratti C. Predicting regional economic indices using big data of individual bank card transactions. arXiv preprint arXiv:150600036. 2015;.

[51] Lane J, Stodden V, Bender S, Nissenbaum H. Privacy, big data, and the public good: Frameworks for engagement. Cambridge University Press; 2014.

[52] Christin D, Reinhardt A, Kanhere SS, Hollick M. A survey on privacy in mobile participatory sensing applications. Journal of systems and software. 2011; 84(11):1928–1946. https://doi.org/10.1016/j.jss.2011.06.073

[53] Bélanger F, Crossler RE. Privacy in the digital age: a review of information privacy research in information systems. MIS quarterly. 2011; 35(4):1017–1042.

[54] Krontiris I, Freiling FC, Dimitriou T. Location privacy in urban sensing networks: research challenges and directions [security and privacy in emerging wireless networks]. IEEE Wireless Communications. 2010; 17(5). https://doi.org/10.1109/MWC.2010.5601955

[55] Walker RM, Andrews R. Local government management and performance: a review of evidence. Journal of Public Administration Research and Theory. 2013; 25(1):101–133. https://doi.org/10.1093/jopart/mut038

[56] Offenhuber D. Infrastructure legibility - a comparative analysis of open311-based citizen feedback systems. Cambridge Journal of Regions, Economy and Society. 2014; 8(1):93–112. https://doi.org/10.1093/cjres/rsu001

[57] O'Brien DT, Offenhuber D, Baldwin-Philippi J, Sands M, Gordon E. Uncharted territoriality in coproduction: The motivations for 311 reporting. Journal of Public Administration Research and Theory. 2017; 27 (2):320–335.

[58] NYC Open Data;. Available from: https://nycopendata.socrata.com.

[59] Boston Data Portal;. Available from: https://data.cityofboston.gov/.

[60] City of Chicago Data Portal;. Available from: https://data.cityofchicago.org/.

[61] Kontokosta C. The price of victory: the impact of the Olympic Games on residential real estate markets. Urban Studies. 2012; 49(5):961–978. https://doi.org/10.1177/0042098011411952

[62] Ester M, Kriegel HP, Sander J, Xu X, others. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Kdd. vol. 96; 1996. p. 226–231.

[63] Hans C. Bayesian lasso regression. Biometrika. 2009; 96(4):835–845. https://doi.org/10.1093/biomet/asp047

[64] Haykin SS, Haykin SS, Haykin SS, Haykin SS. Neural networks and learning machines. vol. 3. Pearson Upper Saddle River, NJ, USA:; 2009.

[65] Girosi F, Jones M, Poggio T. Regularization theory and neural networks architectures. Neural computation. 1995; 7(2):219–269. https://doi.org/10.1162/neco.1995.7.2.219

[66] Jin Y, Okabe T, Sendhoff B. Neural network regularization and ensembling using multi-objective evolutionary algorithms. In: Evolutionary Computation, 2004. CEC2004. Congress on. vol. 1. IEEE; 2004. p. 1–8.

[67] Liaw A, Wiener M, others. Classification and regression by randomForest. R news. 2002; 2(3):18–22.

[68] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research. 2011; 12(Oct):2825–2830.

[69] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Machine learning. 2006; 63(1):3–42. https://doi.org/10.1007/s10994-006-6226-1

[70] Anselin L. Spatial econometrics: methods and models. vol. 4. Springer Science & Business Media; 2013.